PROOF COMPLEXITY AS A COMPUTATIONAL LENS

LECTURE 16: POLYNOMIAL CALCULUS SPACE LOWER BOUNDS

$\boxed{\text{Today}}$: Talk about space in polynomial calculus
Give optimal linear lower bound for
PCR monomial space of refuting
random k-CNF formulas for $k \geq 4$

Polynomial calculus derivation $\boxed{\text{CONFIGURATION-STYLE}}$

$\pi : (\mathbb{P}_0, \mathbb{P}_1, \ldots, \mathbb{P}_\tau)$

$\mathbb{P}_i$ set of polynomials

$\mathbb{P}_0 = \emptyset$

If refutation: $1 \in \mathbb{P}_\tau$

$\mathbb{P}_t$ follows from $\mathbb{P}_{t-1}$ by

(1) AXIOM DOWNLOAD $\quad \mathbb{P}_t = \mathbb{P}_{t-1} \cup \{A\}$ for

- A polynomial in input
- $A \doteq x^2 - x$
- $A \doteq x + \overline{x} - 1 \qquad$ [PCR]

(2) INFERENCE $\quad \mathbb{P}_t = \mathbb{P}_{t-1} \cup \{p\}$

where polynomial p inferred by

$\boxed{\text{LINEAR COMBINATION}}$ $\quad \dfrac{q \qquad r}{\alpha q + \beta r} \quad \alpha, \beta \in \mathbb{F}$

$\boxed{\text{MULTIPLICATION}}$ $\quad \dfrac{q}{mq} \quad m$ monomial

(3) ERASURE $\quad \mathbb{P}_t \subsetneq \mathbb{P}_{t-1}$

$\boxed{\text{MONOMIAL}}$ $m = \prod_i x_i^{a_i}$

Because of Boolean axioms $x_i^2 - x_i$

can assume multilinearity $a_i = 1$

$\boxed{\text{TERM}}$ $t = \alpha m$ $\alpha \in \mathbb{F}$

Polynomial represented as sum of terms =

= linear combination of (distinct) monomials

$$\boxed{p = \sum_{i=1}^{s} \alpha_i m_i}$$

Space measures

$\boxed{\text{LINE SPACE}}$ $LSp(p) = 1$

$\boxed{\text{MONOMIAL SPACE}}$ $MSp(p) = s$

$\boxed{\text{VARIABLE SPACE}}$ $VarSp(p) = \left| \bigcup_{i=1}^{s} Vars(m_i) \right|$

$\boxed{\text{TOTAL SPACE}}$ $TotSp(p) = \sum_{i=1}^{s} Deg(m_i)$

For configuration $\mathbb{P} = \{p_1, \ldots, p_k\}$

$$VarSp(\mathbb{P}) = |Vars(\mathbb{P})|$$

Other space measures M

$$M(\mathbb{P}) = \sum_{p_i \in \mathbb{P}} M(p_i)$$

For all space measures M

$$\pi = (\mathbb{P}_0 = \emptyset, \mathbb{P}_1, \ldots, \mathbb{P}_\tau)$$

$$M(\pi) = \max_i \{ M(\mathbb{P}_i) \}$$

For CNF formula F (or set of polynomials $\mathcal{P}$)

$$M(F \vdash \bot) = \min_{\pi : F \vdash \bot} \{ M(\pi) \}$$

The line space of refuting k-CNF formulas is constant, so we will not talk more about this measure

Monomial space for PCR defined in [ABRW '02] + lower bounds, but only for wide formulas. Turn into 3-CNF ⇒ lower bound breaks

[FLNRT '15]

- PCR monomial space lower for specific 4-CNF formulas (first bounded-width formulas)
- Upper bounds for $\boxed{PC}$ (non-obvious) for k-CNF formulas: Simultaneous exponential size and linear monomial space in PC
- For $\boxed{PC}$ $MSp_{PC}(\widetilde{PHP}^{n+1}_{n} \vdash \bot) = \Omega(n)$
- $\boxed{\text{Weight-constrained CNF formula}}$
  For $C = a_1 \lor a_2 \lor \dots \lor a_w$ w large also have all pairwise clauses $\bar{a}_i \lor \bar{a}_j$

$$\boxed{\begin{array}{l}\text{If } F \text{ weight-constrained, then} \\ Sp_{R}(F \vdash \bot) = Sp_{R}(\tilde{F} \vdash \bot) + O(1) \\ MSp_{PCR}(F \vdash \bot) = \Theta(MSp_{PCR}(\tilde{F} \vdash \bot))\end{array}}$$

[Bonacina-Galesi '15]

Note that FPHP is weight-constrained

- Optimal linear lower bounds for random 4-CNF formulas
- General framework

(but [FLMNV '25] might give nicer exposition of this framework — obviously subjective)

[FLMNV '25] (Conference 2013) INDCO IV

Fix F to be k-CNF formula $k = O(1)$

① $\boxed{MSp_{PCR}(F[\oplus] \vdash \perp) = \Omega(W_R(F \vdash \perp))}$ ANY FIELD

② Take Tseitin formula over expanders, make two copies of each edge in graph $\Rightarrow$
$\exists$ k-CNF formulas F_n of size $\Theta(n)$ s.t.

Still best result for any graphs of degree ≥ 6

$$\boxed{\begin{aligned} Sp_{PC}(F_n \vdash \perp) &= O(n \log n) \\ Deg(F_n \vdash \perp) &= O(1) \\ MSp_{PCR}(F_n \vdash \perp) &= \Omega(n) \end{aligned}}$$

OVER GF(2)

Stronger separation than for resolution!
But depends on characteristic

③ $\Omega(\sqrt{n})$ PCR space lower bounds for Tseitin over RANDOM 4-regular graphs

④ "Barrier result" for BG-framework for FPHP Provably does not work

[Galesi-Kołodziejczyk-Thapen '24] (Conference '19)

$$\boxed{MSp_{PCR}(F \vdash \perp) = \Omega\left(\sqrt{W_R(F \vdash \perp)}\right)}$$

So for all formulas where we have tight width lower bounds, get PCR space lower bound with square root loss

OPEN

- Can clause space and monomial space be separated asymptotically?
- Is monomial space lower-bounded by black-white pebbling price for (substituted) pebbling formulas?
- FPHP monomial space lower bounds?

Why is it hard(er) to prove monomial space lower bounds?

Resolution: For $\pi = (\mathbb{C}_0 = \emptyset, \mathbb{C}_1, \mathbb{C}_2, \ldots, \mathbb{C}_\tau)$

Construct assignment / 1-CNF over disjoint variables

$\mathbb{D}_t$ such that

- $\mathbb{D}_t \vDash \mathbb{C}_t$
- $|\mathbb{D}_t| \leq |\mathbb{C}_t|$

AXIOM DOWNLOAD: Find new variable by expansion

INFERENCE: Do nothing

ERASURE: Just shrink $\mathbb{D}_{t-1}$ by keeping one literal per clause in $\mathbb{C}_t$

This doesn't work for PCR !

Say download

$$\boxed{\begin{array}{c} x_1 - 1 \\ x_2 - 1 \\ \vdots \\ x_n - 1 \end{array}}$$

Space $2n$

Configuration satisfied by assigning n variables

Inference steps ⟿

$$\boxed{x_1 x_2 \cdots x_n - 1}$$

Space 2

Still need to set n variables to satisfy configuration

We cannot afford this !

[ABRW '02] Maintain 2-CNF formula $\mathbb{D}_t$ with clauses over pairwise disjoint sets of variables

Maintain $\mathbb{D}_t \models \mathbb{P}_t$

Stronger inductive hypothesis !

We will see one example of this approach today

RANDOM k-CNF FORMULA F

with n variables and $\boxed{\text{density}}$ Δ

outcome of drawing Δn clauses uniformly at random from set of all k-clauses (exactly k literals) over n variables

$$\boxed{F \sim \mathcal{F}_k^{n,\Delta}}$$

With or without replacement does not really matter, but say with replacement for simplicity

$\boxed{\text{CLAUSE-VARIABLE INCIDENCE GRAPH}}$ G_F

$G_F = (\mathcal{U} \,\dot\cup\, V, E)$

Left vertices = clauses C

Right vertices = variables x

Edge (C, x) if variable x appears in clause C

$G = (\mathcal{U} \,\dot\cup\, V, E)$ is $\boxed{(s, \delta)\text{-BIPARTITE EXPANDER}}$ if $\forall\, \mathcal{U}' \subseteq \mathcal{U}$, $|\mathcal{U}'| \leq s$ it holds that $|N(\mathcal{U}')| \geq \delta \cdot |\mathcal{U}'|$

THEOREM 1 [CS '88]

For any fixed $k \in \mathbb{N}^+$, $k \geq 4$, $\Delta > 0$ there exist $\delta, \gamma > 0$ such that the CVIG G_F of a randomly sampled k-CNF formula $F \sim \mathcal{F}_k^{n,\Delta}$ is a $(2+\delta, \gamma n)$-bipartite expander asymptotically almost surely.

What we want to prove today (almost)

THEOREM 2 [BG'15, BBGHMW'17]
For any fixed $k \in \mathbb{N}^+$, $k \geq 3$ and any $\Delta > 0$ for $F \sim \mathcal{F}_k^{n,\Delta}$ it holds asymptotically almost surely that

$$\boxed{MSp(F \vdash \bot) = \Omega(n)}$$

Remark For any $k \in \mathbb{N}^+$, $F \sim \mathcal{F}_k^{n,\Delta}$ is unsatisfiable a.a.s. for large enough fixed $\Delta = \Delta(k)$.

The proof only uses expansion of G_F
We will prove lower bound only for $k \geq 4$

THEOREM 3
Let F be a CNF formula such that the CVIG G_F is an $(s, 2+\delta)$-bipartite expander for $\delta > 0$ constant. Then $\boxed{MSp(F \vdash \bot) = \Omega(s)}$.

PROOF IDEA

Want to prove $MSp(F \vdash \bot) \geq s$
Show that if $\pi = (\mathbb{P}_0 = \emptyset, \mathbb{P}_1, \ldots, \mathbb{P}_\tau)$ is a PCR derivation in monomial space $\leq s$, then $\mathbb{P}_\tau$ is satisfiable (so $\bot \notin \mathbb{P}_\tau$)

"SHADOW CONFIGURATIONS"

Build sequence of 2-CNF formulas $\mathbb{D}_0, \mathbb{D}_1, \ldots, \mathbb{D}_\tau$ such that:

a) $\mathbb{D}_i$ is satisfiable
b) $\mathbb{D}_i \vDash \mathbb{P}_i$

More concretely, we are going to have

(1) $\boxed{\mathbb{D}_t = Q_t \wedge M_t}$

- clauses in Q_t over pairwise distinct sets of variables
- clauses in M_t over pairwise distinct sets of variables
- For each variable x in a clause $D \in M_t$ there is exactly one clause $C \in Q_t$ containing x
- For each clause $C \in Q_t$, there is exactly one variable x in C that appears in exactly one clause $D \in M_t$

It follows from this:

(i) $|Q_t| = 2|M_t|$ (ii) $Vars(M_t) \subseteq Vars(Q_t)$

(iii) $\mathbb{D}_t$ is satisfiable (satisfy Q_t first, then free literals in all clauses in M_t)

(2) Q_t are subclauses of previously downloaded axioms Q_t = "AXIOM PART"

(3) The variables in M_t are from previously downloaded axioms M_t = "DERIVED PART"

(4) $\boxed{|\mathbb{D}_t| = O(MSp(\mathbb{P}_t))}$

(5) $\boxed{\mathbb{D}_t \vDash \mathbb{P}_t}$

This gives interesting structural information about polynomials derivable in small space from F with expanding CVIG G_F

$\boxed{\text{Any polynomial with } s \text{ monomials derivable in sublinear space from } F \text{ can be forced to } 0 \text{ by assigning at most } O(s) \text{ variables, regardless of } Deg(p).}$

In particular, polynomials like

$$1 - \prod_{i=1}^{n} x_i$$

are not derivable in small space

The proof is by induction over the PCR derivation

Inference steps can be ignored

Need to deal with

- axiom downloads ONLINE MATCHINGS
- erasures LOCALITY LEMMA

(s,r)-ONLINE MATCHING

r-MATCHING on bipartite graph $G=(U \dot\cup V, E)$: Edge set $\mathcal{L} \subseteq E$ such that
- Every $u \in U$ incident to 0 or r edges in M
- Every $v \in V$ incident to 0 or 1 edge in M

$L(\mathcal{L}) = \{u \in U \mid \exists (u,v) \in \mathcal{L}\}$

$R(\mathcal{L}) = \{v \in V \mid \exists (u,v) \in \mathcal{L}\}$

PERFECT r-matching on U' if $L(\ ) = U'$

SIZE of r-matching $= |\mathcal{L}|$ (divisible by r)

$\boxed{(s,r)\text{-ONLINE MATCHING GAME}}$

Start with $\mathcal{L} = \emptyset$. Always $|\mathcal{L}| \leq s$

In each round, Challenger
(a) asks for $u \in U \setminus L(\mathcal{L})$, or
(b) erases $U' \subseteq L(\mathcal{L})$

Matcher responds by
(a) Adding r edges incident to u to $\mathcal{L}$
(b) Remove from $\mathcal{L}$ all edges incident to U'

Challenger wins when Matcher unable to respond
Matcher wins if can play forever

A collection $\mathfrak{L}$ of r-matchings is an $\boxed{(s,r)\text{-ONLINE MATCHING}}$ if
- $\emptyset \in \mathfrak{L}$
- $\mathfrak{L}$ is DOWNWARD CLOSED: If $\mathcal{L} \in \mathfrak{L}$ and $\mathcal{L}' \subseteq \mathcal{L}$ is an r-matching, then $\mathcal{L}' \in \mathfrak{L}$
- If $\mathcal{L} \in \mathfrak{L}$ and $|\mathcal{L}| \leq s-r$, then for any $u \in U \setminus L(\mathcal{L})$ $\mathcal{L}$ can be extended to an r-matching including also u.

THEOREM 4
Let F be a CNF formula such that the CVIG G_F has an $(s,2)$-online matching.
Then $\boxed{MSp_{PCR}(F \vdash \bot) > s/8}$

THEOREM 5 [Feldman, Friedman, & Pippenger '88]
If G is an $(s, r+\delta)$-bipartite expander, then G has an $(s\delta/2, r)$-online matching

Cannot get this for $k=3$ and $r=2$

COROLLARY 6 [Bonacina & Galesi '15]
For every fixed $k \in \mathbb{N}^+$, $k \geq 4$, and for every $\Delta > 0$ fixed, for $\boxed{F \sim \mathcal{F}_k^{n,\Delta}}$ it holds a.a.s. that $\boxed{MSp_{PCR}(F \vdash \bot) = \Omega(n)}$

We will carry out the proof of Theorem 4 by constructing shadow configurations as described before

Example

$$F = \{\bar{a} \lor x \lor \bar{y},\ b \lor \bar{x} \lor y,\ a \lor c \lor z,\ d \lor \bar{z} \lor \bar{w}\}$$

$$Q = \{\bar{a} \lor x,\ b \lor y,\ c \lor z,\ d \lor \bar{w}\}$$

$$M = \{\ a \lor \bar{b}\ ,\ \bar{c} \lor d\ \}$$

To deal with axiom downloads, we can use the $\boxed{\text{online}}$ $\boxed{(s,2)-}$ matching

To deal with erasures we need to prove a LOCALITY LEMMA as done in [ABRW '02] and all later PCR space papers.

LEMMA 7 (LOCALITY LEMMA)

Suppose that $\mathbb{P}$ is a set of polynomials (configuration) and $\mathbb{D} = Q \wedge M$ is a shadow configuration such that $\mathbb{D} \vDash \mathbb{P}$. Then there exists a shadow configuration $\mathbb{D}' = Q' \wedge M'$ such that

- $\mathbb{D}' \vDash \mathbb{P}$
- $Q' \subseteq Q$
- $|M'| \leq 2 \cdot MSp(\mathbb{P})$

Let us see that this is sufficient to prove Thm 4

Proof of Thm 4

Let G_F CVIG for F.

Fix $(s,2)$-online matching M for G_F.

Let $\pi = (P_0 = \emptyset, P_1, \ldots, P_\tau)$ be PCR derivation from F in monomial space $\leq S$

Inductively build sequence of shadow configurations $(\mathbb{D}_0, \mathbb{D}_1, \ldots, \mathbb{D}_\tau)$ so that for ever $\mathbb{D}_t = Q_t \wedge M_t$ it holds that

(a) $\mathbb{D}_t \vDash P_t$

(b) $|M_t| \leq 2 \; MSp(P_t)$

In this construction, we also maintain subsets of axioms $\boxed{A_t \subseteq F}$ such that every $C \in \cancel{A_t} Q_t$ is a subclause of some (uniquely chosen) $D_C \in A_t$ and a perfect 2-matching $\mathcal{L}_t$ from A_t into $Vars(Q_t)$; namely $D_C \in A_t$ gets mapped to $Vars(C)$.

bijection between Q_t & A_t so that this holds

$\boxed{\text{Base case}}$ $\mathbb{P}_0 = \emptyset$

$\mathbb{D}_0 = A_0 = \mathcal{L}_0 = \emptyset$

$\boxed{\text{Inductive step}}$ Case analysis depending on what the derivation does

$\boxed{\text{AXIOM DOWNLOAD}}$ of polynomial p_B encoding axiom clause $B \in F$ or $x^2 - x$ or $x + \overline{x} - 1$

If $\mathbb{D}_{t-1} \vDash p_B$, then set $\mathbb{D}_t = \mathbb{D}_{t-1}$
(In particular, this takes care of Boolean and complementarity axioms)

If $\mathbb{D}_{t-1} \nvDash p_B$ then, in particular, $Q_{t-1} \nvDash p_B$ and so $B \notin A_{t-1}$

By invariant (b) we have that
$|\mathcal{L}_{t-1}| = 2|A_{t-1}| = 2|Q_{t-1}| = 4|M_{t-1}| \leq$
$8\ MSp(\mathbb{P}_{t-1}) \leq 8\ MSp(\mathbb{P}_t) - 8 \leq s - 8$

Therefore, using the $(s, 2)$-~~online~~ matching matching $\mathcal{L}'$ we can extend $\mathcal{L}_{t-1}$ to match also B to two variables x, y not occurring in Q_{t-1} (and hence not in $\mathbb{D}_{t-1}$)

Suppose that $\boxed{B = x^a \vee y^b \vee B'}$
[recall notation x^a = literal over a satisfied by $x \mapsto a$]

We would like to add $x^a \vee y^b \subseteq B$ to Q_{t-1}. But Q_t need to have even size... Annoying...

Add a second dummy axiom download of $B' = p_{B'}$ such that

$$\mathbb{D}_{t-1} \cup \{x^a \vee y^b\} \nvDash p_{B^*}$$

Note that $\mathbb{D}_{t-1} \cup \{x^a \vee y^b\}$ is satisfiable. Therefore, such an axiom B^* exists, or F is satisfiable and there is nothing to prove.

Since $|\mathcal{L}'| \leq s-6$, we can play the online matching game again to match B^* with, say $\{z, w\}$ not in $Vars(Q_{t-1}) \cup \{x, y\}$

Suppose $\boxed{B^* = z^c \vee w^d \vee B^{*\prime}}$

Now we set

$$Q_t = Q_{t-1} \wedge (x^a \vee y^b) \wedge (z^c \vee w^d)$$
$$M_t = M_{t-1} \wedge (x \vee z) \quad \Leftarrow \text{We don't care}$$
$$\mathcal{L}_t = \mathcal{L}_{t-1} + \text{matching } B \lessgtr^{x}_{y} \; B^* \lessgtr^{z}_{w}$$
$$A_t = A_{t-1} \cup \{B, B^*\}$$

This completes the inductive step.

(If we really cared about optimizing constant factors, we could now do a deletion of the fake-downloaded axiom B^*, but we don't need to, since M_t is allowed to grow by 2 clauses at axiom download)

INFERENCE

Just copy everything from time step $t-1$.
The space increased, so we can do this.

ERASURE $\mathbb{P}_t \subsetneq \mathbb{P}_{t-1}$

$\mathbb{D}_{t-1} \vDash \mathbb{P}_t$. If $\mathbb{D}_{t-1}$ is not too large, then copy everything from time step $t-1$.

Otherwise appeal to Locality lemma to get $\mathbb{D}' = Q' \wedge M'$.

Lemma says
$$\mathbb{D}' \vDash \mathbb{P}_t$$
$$|M'| \leq 2 \cdot MSp(\mathbb{P}_t)$$
$$Q' \subseteq Q_{t-1}$$

So set
$$M_t = M' \qquad \mathbb{D}_t = \mathbb{D}'$$
$$Q_t = Q'$$

A_t = axioms corresponding to $Q' \subseteq Q_{t-1}$
$\mathcal{L}_t \subseteq \mathcal{L}_{t-1}$ matching of $A_t \subseteq A_{t-1}$.

This shows that as long as $MSp(\pi) \leq s/8$, we can construct shadow configurations $\mathbb{D}_t \vDash \mathbb{P}_t$. The $\mathbb{D}_t$:s are satisfiable by construction. Hence, so are the $\mathbb{P}_t$:s. $\square$

We need to prove:
- Online matching Theorem 5
- Locality Lemma 7

Thm 5 is standard — used with $r=1$ for resolution clause space lower bounds

Locality lemma is more interesting...

Need the following easy extension of Hall's marriage theorem

LEMMA 8 For $G = (U \dot\cup V, E)$ there is a perfect 2-matching from U into V iff $\forall U' \subseteq U$ it holds that $|N(U')| \geq 2|U|$.

Proof ($\Rightarrow$) By the existence of the 2-matching

($\Leftarrow$) Make 2 copies u', u'' of every vertex $u \in U$

Apply Hall's theorem to get matching

Merge u' & u'' back into u. $\square$

Proof of Locality Lemma

Plan: Identify part $\mathbb{D}_1 \subseteq \mathbb{D}$ that is small enough to keep safely (based on matching with monomials in $\mathbb{P}$)

For $\mathbb{D} \setminus \mathbb{D}_1$ get 2-matching from monomials in $\mathbb{P}$

Pull magic tricks on Q and M

In a bit more detail...

Build bipartite graph $\boxed{H = (U \dot\cup V, E)}$

U = monomials in $\mathbb{P}$ (distinct)

V = clauses in derived part M

Clause $C \in M$ associated to 2 unique clauses in Q, namely those containing the 2 variables in C

For $C \in \mathcal{M}$, denote these clauses $Q(C)$.

Add edge (m, C) in H if $Vars(m) \cap Vars(C) \neq \emptyset$

Let $\boxed{\mathcal{U}_1 \subseteq \mathcal{U}}$ be of maximal size such that $|N(\mathcal{U}_1)| \leq 2|\mathcal{U}_1|$

Let
$$\boxed{\begin{aligned} \mathcal{M}_1 &= N(\mathcal{U}_1) \qquad \mathbb{D}_1 = \mathbb{Q}_1 \cap \mathcal{M}_1 \\ Q_1 &= \bigcup_{C \in \mathcal{M}_1} Q(C) \end{aligned}}$$

If $\mathcal{U}_1 = \mathcal{U}$, then $\mathbb{D}_1 \vDash \mathbb{P}$, since

$Vars(\mathbb{D} \setminus \mathbb{D}_1) \cap Vars(\mathbb{D}) \equiv$

$Vars(\mathbb{D} \setminus \mathbb{D}_1) \cap Vars(\mathbb{P}) \equiv \emptyset$

Otherwise, let

$$\boxed{\begin{aligned} \mathcal{U}_2 &= \mathcal{U} \setminus \mathcal{U}_1 \\ \mathcal{M}_2 &= \mathcal{M} \setminus \mathcal{M}_1 \end{aligned}} \quad \boxed{\begin{aligned} \mathbb{D}_2 &= \mathbb{D} \setminus \mathbb{D}_1 \\ Q_2 &= Q \setminus Q_1 \end{aligned}}$$

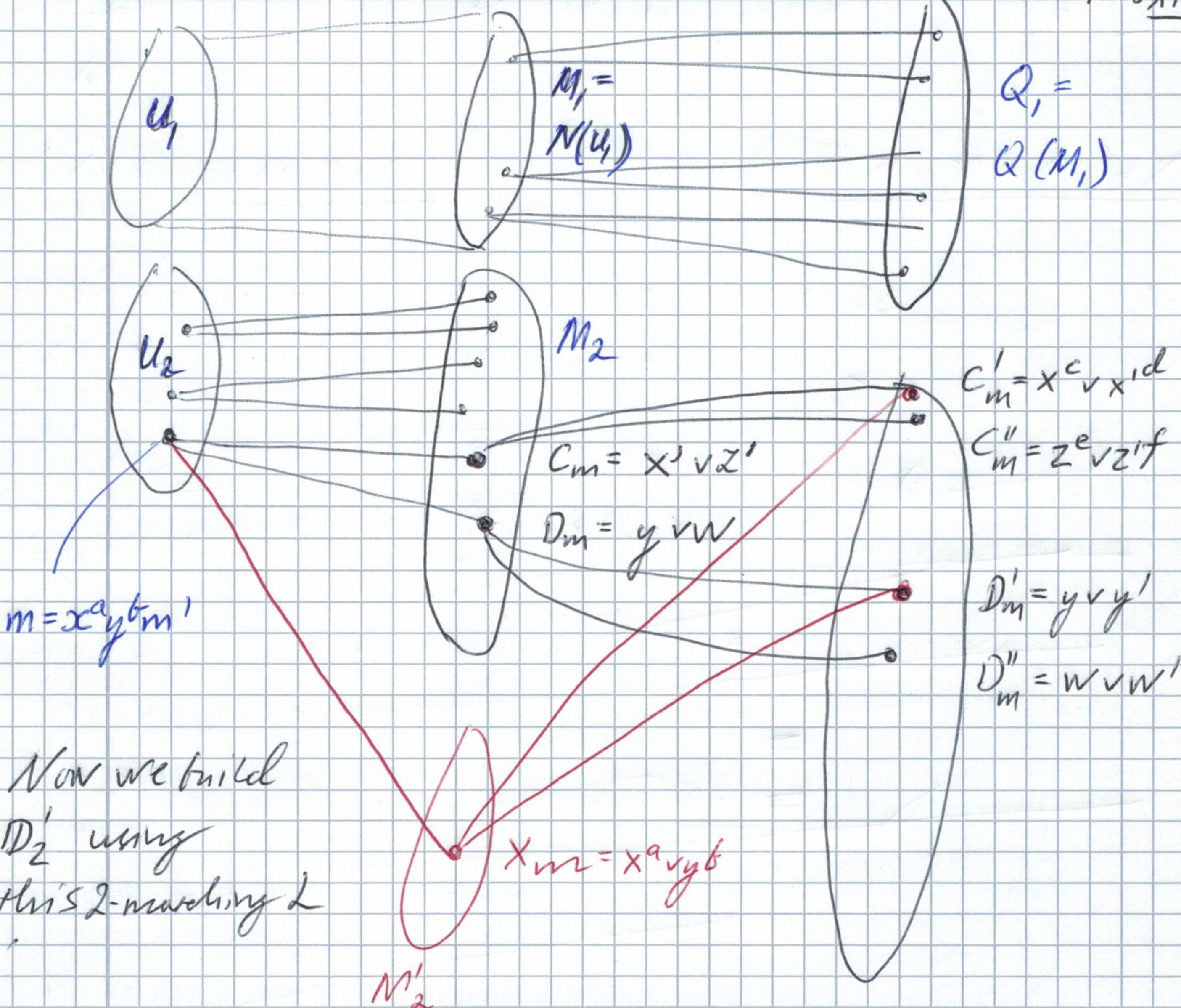
For any $\mathcal{U}' \subseteq \mathcal{U}_2$ we have

$$|N(\mathcal{U}') \cap \mathcal{M}_2| > 2|\mathcal{U}'|$$

since otherwise $\mathcal{U}'$ could be added to $\mathcal{U}_1$ to get larger set

So there is a 2-matching from $\mathcal{U}_2$ into $\mathcal{M}_2$.

(In partical, any constant monomial is in $\mathcal{U}_1$)



Now we build D'_2 using this 2-matching L

Fix $m \in U_2$ matched to C_m & D_m

$Q(C_m) = \{C'_m, C''_m\}$

$Q(D_m) = \{D'_m, D''_m\}$

as in figure

$\mathrm{Vars}(m) \cap \mathrm{Vars}(C'_m, C''_m) \neq \emptyset$

$\mathrm{Vars}(m) \cap \mathrm{Vars}(D'_m, D''_m) \neq \emptyset$

By construction

Say
$$\boxed{\begin{array}{l} m = x^a y^b m' \\ x \in \mathrm{Vars}(C'_m) \\ y \in \mathrm{Vars}(D'_m) \end{array}}$$

Construct new clause

$$\boxed{X_m = x^a \vee y^b}$$

Note that any assignment satisfying X_m will zero out m

X_m is added to M'

Clauses from Q_2 containing x & y are added to Q' and associated to X_m

In our case

$$Q(X_m) = \{C_m', D_m'\}$$

Do this for every $m \in U_2$

Then set

$$\boxed{\begin{aligned} M_2' &= \bigwedge_{m \in U_2} X_m \\ Q_2' &= \bigwedge_{m \in U_2} Q(X_m) \\ \mathbb{D}_2' &= Q_2' \wedge M_2' \end{aligned}}$$

$$\boxed{\begin{aligned} M' &= M_1 \wedge M_2' \\ Q' &= Q_1 \wedge Q_2' \end{aligned}}$$

Finally $\boxed{\mathbb{D}' = D_1 \wedge \mathbb{D}_2'}$

Clearly $Q' \subseteq Q$.

By construction

$$|M'| = |M_1| + |M_2'| \le 2|U_1| + |U_2| \le 2\, MSp(P_t)$$

But we need to prove $\boxed{\mathbb{D}' \vDash \mathbb{P}}$!

Show this in somewhat roundabout way

Given any α satisfying $\mathbb{D}'$
Build assignment β that
- satisfies $\mathbb{D}$
- agrees with α on all monomials in $\mathbb{P}$

$\beta(\mathbb{P}) = T$ since $\mathbb{D} \models \mathbb{P}$

But then also $\alpha(\mathbb{P}) = T$, since α & β agree on all monomials.
Since α is arbitrary, we conclude

$$\mathbb{D}' \models \mathbb{P}$$

Given α satisfying $\mathbb{D}'$
Define β_1 on $Vars(\mathbb{D}')$ to agree with α

Now construct β_2 on $Vars(\mathbb{D}) \setminus Vars(\mathbb{D}')$

Fix any clause $C \in M_2$, $C = x^g \vee z^h$

Let $Q(C) = \{ x^c \vee x'^d, \quad = C'$
$\qquad\qquad\quad z^e \vee z'^f \} \quad = C''$

Case 1: C matched with $m \in U_2$

$|Vars(C) \cap Vars(X_m)| = 1$ say $x \in Vars(X_m)$

$|\{C', C''\} \cap Q'| = 1$ say $C' \in Q'$

Then C' satisfied by α — want to keep that, but already taken care of, since $C' \in \mathbb{D}'$

In β_2 set $\beta_2(z) = h$
$\beta_2(z') = f$

Case 2: C not matched EASY

Then we have no restriction

Let, e.g.,

$$\beta_2(x) = g \qquad \beta_2(x') = d \qquad \beta_2(z') = f$$

to satisfy C and $Q(C)$

After this case analysis, fix all remaining variables in $Vars(\mathbb{D}) \setminus Vars(\mathbb{D}')$ so that β_2 agrees with α

Set $\boxed{\beta = \beta_1 \cup \beta_2}$ assign disjoint sets of variables

Now:

- β_2 satisfies $\mathbb{D} \setminus \mathbb{D}'$ by construction
- β_1 satisfies $\mathbb{D} \cap \mathbb{D}' = \mathbb{D}_1$ since it agrees with α there

$Vars(\mathcal{U}_1) \subseteq Vars(\mathbb{D}_1)$ so α & β agree on monomials $m \in \mathcal{U}_1$

α and β_2 both satisfy $\mathcal{U}_2'$,

so $\alpha(m) = \beta_2(m) = 0$ for $m \in \mathcal{U}_2$

This concludes the proof. $\square$